State Lottery Revenue: The Importance of Game Characteristics

State Lottery Revenue: The Importance of Game Characteristics[*]

Thomas A. Garrett
Federal Reserve Bank of St. Louis
Research Division
411 Locust Street
St. Louis, MO 63102
tom.a.garrett@stls.frb.org
(314) 444-8601

Russell S. Sobel
Department of Economics
P.O Box 6025
West Virginia University
Morgantown, WV 26506

Abstract

Previous studies find state lottery sales are significantly influenced by socioeconomic characteristics of the population. We extend this literature by examining how the overall expected value, the top prize, and the total combinations influence sales after controlling for these other socioeconomic factors. We perform our empirical analysis on an unparalleled set of data that includes information for 135 on-line lottery games in the United States. Our results show that sales are significantly influenced the top prize amount and odds of winning it, but that sales are not significantly affected by the expected value of the remaining lower prizes.

State Lottery Revenue: The Importance of Game Characteristics

I. Introduction

Thirty-seven states and the District of Columbia now offer state-sponsored lotteries. Although state lotteries provide entertainment for millions of people nationwide, the primary goal of a lottery is to maximize lottery revenues for the state. According to Clotfelter and Cook (1989, page 11), "Lottery agencies are not merely acting out of a liberal respect for consumer sovereignty. They are engaged in a well-focused quest for increased revenues." In fact, the operating papers governing most state lotteries generally cites revenue maximization as the goal of the bureaucracy charged with administering the lottery.[1] In 1999, 34 percent of total lottery ticket sales, or nearly $12 billion, was generated in net lottery revenue to support state government spending in the United States.[2]

A fairly sizeable body of literature has explored the relationship between the socioeconomic characteristics of a state's population and lottery ticket sales.[3] These studies generally find that lottery sales are higher for individuals who belong to minority groups,

---

[1] The stated mission for each state lottery can be found on their website or annual report. Several studies have examined political issues surrounding state lottery adoption. See Davis, Filer and Moak (1992), Alm, McKee and Skidmore (1993) and Garrett (1999).

[2] Net lottery revenues are total lottery sales minus prize payouts, commissions, and administrative expenses. The main reported uses of this revenue were education (51.6%), states' general funds (24.2%), and local programs and infrastructure (13.3%). These data are potentially misleading, however because of fungibility issues (e.g. flypaper effect). See Mikesell and Zorn (1986) for a discussion and evidence. Even if the revenue is earmarked toward education, for example, the state can reduce its own general fund spending on education, effectively converting the increased earmarked revenue to money in the general fund.

[3] See Stover (1987), Clotfelter and Cook (1989), Mikesell (1989), and Hansen (1995) for several representative examples of research in this area.

individuals with little or no formal education, residents of urban areas, and individuals between the ages of 45 and 65. Of all socioeconomic characteristics studied, measures of income have received the most attention. The coefficient on the income variable lends insight into whether the revenue burden of the state lottery is regressive, proportional, or progressive. Most studies have found that overall lottery expenditures tend to be regressive, meaning lower income groups bear a higher burden of the lottery tax as a percent of their income.[4]

These previous studies have provided insight into how the differences across state lottery sales can be explained by socioeconomic characteristic of state residents. We build upon this previous literature by exploring how lottery game sales also depend on game characteristics such as the value of the top prize, the chances of winning, and the overall expected value of the lottery. A few previous studies have examined the importance of game characteristics on predicting lottery revenues for a given state lottery. The first of these studies was conducted by Vrooman (1976). Focusing on the New York Lottery, Vrooman included the expected value of the lottery and the probability of winning the top prize, along with other game characteristics specific to the New York Lottery, in a regression to predict lottery sales. Vrooman found that both the expected value of the lottery and the probability of winning the top prize were not significant determinants of lottery ticket expenditures in New York. In a later study of the Washington State Lottery, Thiel (1991) found that worsening the odds of winning the jackpot led to an increase in lottery revenues, as more players were attracted to the larger jackpots that result from fewer lottery winners. Similarly, Scoggins (1995) examined the Florida Lotto and found

---

[4] The literature on the tax incidence of state lotteries is quite extensive. See Clotfelter (1979), Clotfelter and Cook (1987), Mikesell (1989), Scott and Garen (1994), Hansen (1995), and Stranahan and Borg (1998).

that lottery expenditures, and thus net revenues to the state, will be increased by allocating a

larger percentage of sales to the top prize in the Florida Lotto. The significance of game

characteristics in these studies of individual lotteries supports our conjecture that differences in

game characteristics should also be important in explaining differences in lottery sales across

games.

Our objective is to examine how the prize and odds structure of different lottery games

impacts lottery game sales. In accordance with the stated mission of state lottery agencies,

several studies have shown that state lottery agencies do indeed structure their lottery games to

maximize revenue.[5] In this study, we wish to see which prize and odds structure across states

and games provides the greatest revenue to state lottery agencies. We use a set of unparalleled

data on prize payouts, odds of winnings, and lottery game sales for 135 on-line lottery games in

the United States in 1995 to develop several measures of game characteristics to include in our

empirical analysis.[6] Our model is estimated on the full sample of lottery games, and also on

subgroups of games by type (Lotto versus smaller prize games, for example). Our results show

that the amount of the top prize and the odds of winning it are significant in explaining

differences in sales across U.S. lottery games. However, after controlling for these

characteristics of the top prize, the overall expected value seems to have no significant,

independent influence on lottery sales. We find that the explanatory power of our models

including these game characteristics far exceeds the explanatory power achieved by using only

---

[5] See Mason, Steagall, and Fabritius (1997) and Garrett (2002).

[6] On-line games are games for which a player generally gets tickets from a computer
terminal after filling out a playslip. These are in contrast to instant or scratch-off lottery games,
which are not considered here.

socioeconomic characteristics, as has been done in previous studies of lottery sales. Finally, we find support for the out-of-sample predictive power of our model using a recent change to the Multi-state PowerBall game.

II. The Expected Value and Expected Payout of a Lottery Ticket

The expected value of purchasing a lottery ticket is negative to the lottery player because the expected prize payout on a ticket is less than the dollar cost of purchasing the ticket. The expected value of purchasing a lottery ticket, however, can differ substantially across lottery games because of differences in the expected prize payout across different lottery games. In this section we derive expressions that allow us to calculate the expected prize payout on a given lottery ticket from the general prize structure of each game.

Although the computation of expected prize payout for a lottery ticket is a basic summation of each possible prize amount multiplied by the probability of the prize, the presence of parimutuel lottery prizes requires a modification of the simple formula because the prize amounts are not fixed. This section discusses parimutuel and fixed prizes and presents the derivations and final expression for our expected prize payout variable that considers both fixed and parimutuel prize lottery games.

The state determines the structure of the lottery by setting the prize amounts, number of combinations, and the probabilities of winning each prize. There are two general ways in which prize amounts may be set by the state. The first type is a "fixed" prize for which the prize amount won by the player is a certain fixed amount that is independent of the number of other

players winning that prize and of the outcomes on other prizes. The second type of prize is a "parimutuel" prize for which the prize amount won by the player depends on how many other players win that same prize and the total dollar amount wagered for the lottery drawing. For a parimutuel prize lottery, the state allocates a fixed percentage of ticket sales to each prize level (or prize "tier"), then splits this total amount among all winners of that prize. Some lottery games (like Daily Number games) have all fixed prize amounts, but most large Lotto games have parimutuel prize amounts for the largest several prizes and fixed prize amounts for the lower prizes.

Because lottery games can be comprised of parimutuel prize amounts, the question exists as how to determine the expected prize payout of a lottery game ticket for a lottery that includes parimutuel prizes, since the prize amounts are not a fixed dollar amounts. Consider an all parimutuel lottery game that offers three prize tiers, with 25 percent of total sales allocated to pay jackpot winners, 15 percent of total sales allocated to pay second prize winners, and 10 percent of sales allocated to pay third prize winners. From this example it is clear that the state will allocate 50 percent of total sales to prizes. If the state allocates 50 percent of total sales to prize amounts, then by definition 50 percent of total sales will be given away as prizes. Intuitively, it can be seen that across all players, half of the money wagered is returned as winnings, and thus the average player will have an expected prize amount equal to 50 cents per dollar wagered. We now turn to a more formal derivation of the fact that the expected prize payout value of this lottery game is simply the share of sales allocated to prizes.

From an aggregate perspective, the state lottery's expected total prize payout (*TP*) to all players on a given drawing is by definition equal to the sum across all *N* tickets sold of the expected prize payout (*EP*) for each ticket (*n*).

$$TP = \sum_{n=1}^{N} EP_n \qquad\qquad \textbf{(1)}$$

Because ex ante all tickets have equal expected prize payout values (and thus $EP_n = EP$ for all *N* tickets), the right hand side of the equation simply reduces to the expected total payout being equal to the number of tickets sold, *N*, times the expected prize payout for a single ticket (*EP*).[7] Given that prizes are all won, in the case of an all parimutuel lottery game the total prize payout for the state is not random at all, but is a fixed share of the amount wagered on the drawing. However, it is possible on any given drawing for a parimutuel lottery that there is no winner in one or more of the prize tiers. In most cases this money is distributed across other prizes, but for unwon top prizes, the state lottery generally rolls this amount over to the next drawing. This does cause the expected total prize payout to vary slightly from drawing to drawing on the occasion when it occurs. However, the *average* expected payout value of a ticket across all drawings that year (both the ones in which there was no winner and the drawing with the

---

[7] Here we assume for mathematical simplicity that we are considering a randomly selected ticket. If players tend to clump together on certain number combinations, then the players on these combinations will have below average expected values, while players on sparsely chosen numbers will have above average expected values. Taking this into account (by leaving the summation in the equation, but dividing both sides by *N*) simply yields that what we derive is the *average* expected prize payout value of a ticket, which is precisely our goal.

eventual winner) remains unaffected.[8] Thus, more properly the variable we are deriving here in this section may be defined as the average expected payout value for the lottery. In our later empirical work we incorporate both this average value and a measure of its variance.

Continuing with the derivation, for mathematical simplicity we assume each ticket costs one dollar, and that N ticket are sold. Letting S denote the share of the lottery revenue for the drawing that is allocated to the prize pools (for all prizes), then the state's total payout for a drawing, the left hand side of the equation) will simply be equal to $N{\cdot}S$. Thus equation (1) may be rewritten as:

$$S{\cdot}N = N{\cdot}EP \qquad\qquad\qquad (2)$$

After dividing both sides by N, the intuitive result that the expected prize value of each ticket (EP) is simply the share of the ticket revenue devoted to prizes is easily shown.

$$S = EP \qquad\qquad\qquad (3)$$

Thus, for an entirely parimutuel lottery game, the expected prize payout value of a given ticket is simply the share of ticket revenue devoted to the prize pools, a result that was intuitively straightforward. When a lottery game consists of some prizes that are fixed and some prizes that are parimutuel, the total expected prize payout value is simply the sum of the expected prize

---

[8] Assuming that the jackpot is won in the next drawing, then the average expected prize amount across the two drawings is equal to the share of revenue allocated to prizes. Because we use annual data in our later empirical analysis, this simplifying assumption is valid to aid us in deriving the average expected prize payout value on the game over the year. However, the expected value of a ticket can vary slightly from drawing to drawing in the event of a rollover. In our empirical analysis we incorporate a variable to control for the impact of the variance in the expected prize payout value over the year.

payout value of the fixed prize games and the expected prize payout value of the parimutuel

prize games.

The expected prize payout value of a ticket for an all fixed prize lottery game is given by:

$$EP = \sum_{k=1}^{K} P_k \cdot W_k = \sum_{k=1}^{K} \left(\frac{c_k}{C} \right) \cdot W_k \tag{4}$$

where k represents one of K total prizes, P_k is the probability of winning prize k, and W_k is the

dollar amount of prize k. The probability of winning prize k is found by dividing the number of

combinations that win prize k, c_k, by the total number of combinations, C, available to the

lottery player, thus $P_k = c_k/C$.

Most large lottery games consist of some prizes that are parimutuel (generally the top

prizes) and some prizes that are fixed. In such cases, the expected prize payout value of a ticket

is simply the sum of the expected prize value across the fixed prizes plus the expected prize

value across the parimutuel prizes.

$$EP = EP_{fixed} + EP_{parimutuel} \tag{5}$$

Now, substituting the results of equations (3) and (4) into equation (5) gives:

$$EP = EP_{fixed} + EP_{parimutuel} = \sum_{fixed} P_k \cdot W_k + S = \sum_{fixed} \frac{c_k}{C} \cdot W_k + S \tag{6}$$

8

where, for lottery games comprising of some or all parimutuel prizes, S again represents the total percent of ticket revenue devoted to all parimutuel prizes.[9] For a lottery game offering only fixed prize amounts, computing the expected prize payout value remains the same as in equation (4), because S would equal 0 in equation (6) for such a lottery. Lottery games having only parimutuel prizes will have an expected prize payout value equal to S, the same as the result in equation (3), and finally those lottery games having both fixed and parimutuel prizes will have an expected prize payout value equal to the expected prize payout value of the fixed prize amounts plus the total percent of ticket sales devoted to the parimutuel prize amounts.

United State's Lottery Games

The expected prize payout value for our sampling of lottery games is computed using data on lottery game sales and prize and odds structures for 135 Lotto, Daily Number, and nightly Keno games available in the United States during 1995.[10] Here we briefly describe each type of game used in our analysis.

---

[9] An important aspect of this equation is that the expected prize payout values computed are essentially the true expected prize payout values, as they essentially use the net present value of discounted jackpot prizes instead of using the advertised nominal jackpot amount. For most on-line lottery games, the share of lottery sales earmarked for the jackpot is used to buy an annuity, which pays the advertised jackpot amount over many years. The advertised jackpot amount is thus much higher, and using this figure to compute the expected prize payout value would result in overestimation. Our use of the share amounts that are used to buy the annuities circumvents this problem.

[10] Total 1995 annual sales for each lottery game are from Public Gaming International's *Public Gaming Magazine*, September 1996. Prize and odds information and the frequency of offering for each game was obtained by contacting each state lottery agency.

Lotto games typically offer parimutuel, multi-million dollar jackpots, with odds of winning the jackpot set at several million to one. These games require that players match 5 of 5 or 6 of 6 numbers out of a given range of numbers in order to win the top prize, which is usually a parimutuel prize amount. If a player's numbers match those drawn, he or she wins the top prize (or a portion of a parimutuel top prize if there is more than one winner). Lottos have the added feature of a "roll-over," meaning that if there is no jackpot winner, the prize money allocated to the jackpot is added to the jackpot prize money for the next drawing. Some Lotto games are available in numerous states. These multi-state games involve several states participating in the same lottery game. Due to the administrative costs of lottery games, states having a low population, and thus on average low sales, cannot offer multi-million dollar jackpots. But if several states pool their sales, these large jackpots can be offered.[11] The multi-state games available in 1995 included PowerBall, Tri-West Lotto, Tri-State Card Cash and Tri-State Megabucks.

Daily Number games, sometimes called three-digit and four-digit games, are offered six or seven nights a week. For the three-digit game, a player picks a number between 000 and 999, and wins a top fixed prize of $500 (in most states) if his or her three-digit number matches the three-digit number drawn. Similarly, for the four-digit game, a player picks a number between 0000 and 9999, and wins a top fixed prize of $5,000 if his or her four-digit number matches the four-digit number drawn.

The final type of lottery game, Keno, a fixed prize game, made its debut into the lottery portfolio of many states several years ago. The "Keno computer" randomly chooses 20 numbers

---

[11] Deboer (1985) discusses economies of scale and administrative costs of state lotteries.

between 1 and 80. The player can choose to play any one of ten number games, and hopes to

match numbers drawn by the computer. For example, if a player chooses to play the ten number

game he selects 10 numbers between 1 and 80 on the playslip. Twenty numbers between 1 and

80 are then randomly drawn and displayed on a monitor. If some or all of the 10 numbers

selected by the player match the numbers drawn by the computer, the player wins. In some

states Keno can be played at bars, restaurants, and even gas stations, and is offered every 5

minutes between the hours of roughly 6:00am and 2:00am on video monitors.   Other states offer

a single Keno drawing several times per week.[12]

III. Empirical Methodology and Data

The expected prize payout value of a lottery incorporates information about every prize

amount and the odds of winning each prize. However, a substantial portion of a lottery's

expected prize payout value generally derives from the very large top prize, despite the usually

remote odds of winning it. It is these top prize amounts that are generally used to advertise the

lottery and it is these amounts that are the material for news stories when rollovers make them

grow to unusually large amounts. Because of the potential for the top prize component to have a

distinct influence on sales relative to the value from the lower prizes, we also include the top

prize amount and the total number of combinations as individual variables in our regression

---

[12] For our analysis, we only consider Keno games that are offered nightly. For those
states that offer Keno every five minutes, data is not available on sales per drawing per number
game.   Sales for each game could have been computed by dividing total sales by the number of
games, under the assumption that sales are equal for each Keno game. Given the wide variance
in top prize offerings for Keno games ($2.00 to $100,000), it is very unlikely that all Keno
games experience the same level of sales.

analysis. Although the three game characteristic variables are our main focus, we also include

several standard demographic variables that have been shown to be important determinants of

lottery sales in other studies. We want to include these variables to control for other

characteristics that may impact lottery sales, and to ensure that our game characteristic variables

are not capturing the impact of any omitted variables. In addition, we include the number of

lottery retail outlets available in each state to control for the availability of lottery tickets.

Our analysis involves an estimation of the following model of lottery sales using data

from on-line lottery games available during 1995:

$$PC\ Sales_i = \beta_0 + \beta_1 \cdot (1/EP_i) + \beta_2 \cdot Top\ Prize_i + \beta_3 \cdot Total\ Combinations_i + \beta_4 \cdot Variance_i + \Gamma_i + \varepsilon_i$$

$PC\ Sales_i$ is per capita lottery sales for game i, $Top\ Prize_i$ is the top prize for game i, $Total$

$Combinations_i$ is the total number of combinations for game i, and Γ is a vector of demographic

variables corresponding to lottery game i's state. Since we are estimating an inverse demand

curve for lottery games, the price of a lottery ticket can be represented by $1/EP_i$, which is the

reciprocal of the expected prize payout value of lottery game i.[13] It is very important to note that

the inclusion of the top prize and the number of combinations as separate variables affects the

interpretation of the coefficient on EP_i. The coefficient now reveals the marginal impact of

increasing the expected prize payout value of the lottery, but holding constant the value and odds

_____

[13] The price of a lottery ticket can also be represented as 1-EP. Both specifications
provide similar results.

12

of winning the top prize. This coefficient essentially is a reflection of how an increase in the

expected prize payout value from the lower prize tiers affects lottery sales. The coefficients on

Top Prize$_i$ and *Total Combinations$_i$* will reveal how players respond to changes in the prize or

odds structure of the top prize lottery game holding constant the overall expected prize payout

value of the lottery (a reallocation of the prize value toward the jackpot and away from the

lower prizes). Based on the findings of Thiel (1991) and Scoggins (1995), we expect a positive

coefficient on *Top Prize$_i$* and a negative coefficient on *Total Combinations$_i$*. Because only one

combination wins the jackpot, the odds of winning the top prize decline with a larger number of

combinations. We expect an increase in the total number of combinations to decrease lottery

game sales.[14]

Several studies using time-series game level data have shown that the true payout rate of

a lottery may not be equal to its expected prize payout value (true long-run payout value)

because of jackpot roll-overs and variability in lottery sales over time.[15] In an attempt to capture

the impact of roll-overs and how they may cause changes in expected prize payout values in our

cross sectional data we consider the variance of prize outcomes for each lottery game *i*. Those

lottery games having a greater variance of prize outcomes are more likely to experience jackpot

---

[14] Note that depending on the magnitudes of each coefficient, restructuring a lottery
game's top prize and number of combinations could lead to an increase or decrease in the
expected return, depending upon the magnitude of each coefficient. We also note for reference
that the raw correlation between the top prize and the total number of combinations is 0.407.

[15] See Gulley and Scott (1993), Scott and Gulley (1995), and Mason, Steagall, and
Fabritius (1997)

rollovers because the odds of winning the top prizes become more remote. Let σ_v^2 represent the variance of prize outcomes to an individual player. It is given by:[16]

$$\sigma_v^2 = \sum_{k=0}^{K} (W_k - EP)^2 \cdot P_k = \sum_{k=1}^{K} W_k^2 \cdot P_k - EP^2 \tag{7}$$

The first part of equation (7) shows the standard formula for computing the variance of a random variable. Here it is the sum over all prize outcomes (including the losing outcome of $k=0$) of the squared difference between the prize amount and the expected prize payout value of the lottery. This is further simplified to the sum of the prize amounts squared times their associated probabilities minus the squared expected prize payout value of the lottery. Because the losing prize of zero disappears in this version, it is shown as the summation over only the winning prizes ($k=1$ to K instead of $k=0$ to K). Again recognizing the distinction between fixed and parimutuel prizes, the general formula for the variance, which considers equation (6) for the expected prize amount of a single prize, k, can be represented as

$$\sigma_v^2 = \sum_{fixed} W_k^2 \cdot P_k + \sum_{parm} \frac{s_k^2}{P_k} - EP^2 \tag{8}$$

where the summations in the equation are shown separately for the fixed and parimutuel prizes, s_k represents the share of the ticket sales allocated to the prize pool for a parimutuel prize k, and as before, the probability of winning prize k, P_k, is simply c_k / C.

Descriptive statistics for all variables used in the analysis are presented in Table 1.

[Table 1 about here]

---

[16] This equation uses the standard simplification: $\sigma^2 = E[X - E[X])^2] = E[X^2] - (E[X])^2$.

14

IV. Empirical Results

We provide the results from three regression specifications. The first regression consists of all lottery games in our sample. The second regression considers only the subset of Lotto games, and the third regression considers only the subset of Daily Number and smaller prize lottery games. In addition, to examine the improvement from the previous literature's use of only socioeconomic characteristics, each regression was also run excluding the three game characteristic variables.

Because we are including the top prize of game i as an independent variable, the potential problem of simultaneity arises between game sales and the top prize for those games having a parimutuel top prize. For each drawing (and especially in the case of jackpot rollovers), announcement of the jackpot will induce lottery sales, and these sales will in turn impact the top prize since the top prize is a function of lottery sales. To ensure consistent coefficient estimates, we explored whether the top prize variable should be made endogenous in each regression model. We performed a Hausman test for the following hypothesis: H_0: a single equation OLS model is appropriate because there is no correlation between the top prize variable and the equation error term (top prize is exogenous); or H_1: a system of equations is required due to correlation between the top prize variable and the error term (top prize is endogenous). The details on this test can be found in Hausman (1978) and Godfrey (1988, chapter 5). For each regression model the computed Hausman statistic (shown in Table 2) was less than conventional critical levels from a χ^2 distribution, thus suggesting no significant endogeneity of the top prize variable and that a single equation regression is appropriate.

15

Our regression results are shown in Table 2.

[Table 2 about here]

In none of the three regressions is the coefficient on expected prize payout value significant. Recall that this must be interpreted as the impact of a change in the expected prize payout value holding constant the top prize and number of combinations. Thus, this result suggests that marginal differences in the value of the lower prize amounts have no significant explanatory power in explaining differences in per capita lottery sales. This is also supported by the insignificant coefficient on variance. The implication is that lottery players appear to be fairly unresponsive to changes in the structure of the expected prize payout value from lower prize tiers. The coefficients on the top prize and total combinations show, in support of the findings from Thiel (1991) and Scoggins (1995), that large Lotto games having a higher top prize are likely to have higher sales. Also, an increase in the total combinations available in a Lotto game, thus a worsening of the odds of winning, results in lower lottery sales. The coefficient from the Lotto regression suggests that a $100,000 increase in the top prize will lead to an $1.58 increase in per capita sales, while increasing the total number of combinations by one million will lead to a 51 cent decrease in per capita lottery sales. These results for the Lotto games, reported in column (4) of the table, hold in the full sample regression, but do not hold individually for the smaller prize games. This suggests that the top prize amounts are more influential in determining lottery sales for big jackpot games than for smaller prize games.

We also find support for the previous findings that nonwhite population and income are significant determinants of lottery sales. However, several of the other demographic control variables are insignificant. For the full sample and Lotto subsample regressions, the F-test

16

shows the inclusion of the game characteristic variables significantly increases the explanatory value of the model. The difference, particularly for the subsample of Lotto games, is striking. The percent of the cross game variation in per capita sales explained by the model for Lotto games jumps from about 15 percent to over 52 percent with the inclusion of these three variables. It is clear that a substantial portion of the variation in U.S. Lotto game sales can be explained by the prize structure of the games. It also points to possible, severe omitted variable biases in previous studies of Lotto expenditures that exclude any information about the game characteristics in the models. For the smaller prize games, however, there is not much improvement in explanatory power from the game characteristic variables, and socioeconomic characteristics are significantly greater predictors of smaller game sales than they are of Lotto sales. The significant and substantial differences between the estimated regressions for Lotto games and for the smaller games suggests that these are fundamentally different subsamples of lottery sales revenue. This also has the important implication that other empirical research on state lotteries should avoid techniques that pool these games together.

While our analysis is mostly concerned with the impact of the game characteristic variables, it is worth noting that the estimated magnitude of several socioeconomic variables differs between the Lotto and smaller prize game regressions. Most notably, the income coefficient and the coefficient on nonwhite population are both larger for smaller prize games. While further exploration would be necessary, this might appear to have potentially important implications for incidence analysis of different types of lottery games.

A 'Real-World' Test of Our Model - Evidence from PowerBall

In this section we provide evidence of the predictive power of our Lotto model from column (2) of Table 2. In 1997, the Multistate Lottery Association worsened the odds of winning the PowerBall jackpot from roughly 55 million to one to 80 million to one in an attempt to generate additional revenues through higher jackpots. How accurate is our Lotto model in predicting PowerBall sales, both before and after the prize and odds change? The results are shown in Table 3.

[Table 3 about here]

Before the change in game structure, actual PowerBall sales per capita in 1995 were $25.17. Using the expected prize payout value, variance, top prize and total combinations for PowerBall in 1995, we predict 1995 per capita PowerBall sales of $27.16 using model (4) in Table 2. A year after the prize structure changed, actual 1998 per capita PowerBall sales increased to $30.67. Using the new values for expected prize payout value, variance, top prize and total combinations, our model predicts an increase in per capita PowerBall sales in 1998 to $29.76.[17] Our model thus predicts quite well the real-world behavior of PowerBall sales both before and after the prize and odds structure was changed, and further reveals the importance of considering lottery game characteristics in any model of lottery sales.

V. Policy Implications and Concluding Remarks

---

[17] The variance of Powerball was 180,844 before the change in game structure and 317,343 after the change in structure.

As the demand for state government services continues to increase, states are becoming pressured to find alternative sources of revenues. Many states have adopted lotteries in hopes of easing their fiscal pressures. Given the billions of dollars in lottery revenue generated each year, previous scholarly works have examined which individuals, in terms of many socioeconomic characteristics, are more likely to play the lottery. This information potentially allows states to target certain socioeconomic groups in an attempt to further increase lottery revenues.

Previous works that have examined the socioeconomic determinants of lottery ticket expenditures have overlooked one key aspect of lottery play, namely the importance of lottery game characteristics. Using data from a sampling of on-line lottery games in the United States, we included the expected prize payout value, top prize, and total combinations from each game in a model of lottery sales. We find that game characteristics for Lotto games are significant determinants of lottery sales. Our paper is the first to distinguish between top prize and lower prize tiers in order to see how players react to changes in either. We find that while lower prize tiers may be appropriate in terms of offering a 'consolation' prize, the expected prize payout value of these lower prizes does not seem to be a significant factor in explaining cross game variations in lottery sales. The results from our sampling of U.S. lottery games supports the findings of single game studies that have shown that large jackpots stimulate lottery play. Support for our model was further validated with a real-word test using PowerBall.

Lotteries will undoubtedly continue into the future. But as lotteries face competition from increasing casino and Internet gambling and interstate tax competition, states who are interested in keeping their lotteries viable will have to devise a strategy to generate the most revenue possible. By considering the lottery game characteristics presented in this paper, in

19

addition to the socioeconomic makeup of residents, states can better predict and structure their lottery games in an attempt to maximize revenues.

Finally, our paper has implications for future (and previous) empirical work on state lotteries. Our results not only suggest that empirical studies should be including measures of game characteristics to avoid omitted variable bias, but also that substantial differences between large Lotto games and other types of games necessitate that empirical work consider data from these games separately.

Table 1 - Descriptive Statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
Per Capita Sales	21.117	22.999	0.141	157.961
Expected Prize Payout Value	0.502	0.041	0.241	0.661
Variance	97,893	280,411	243.44	2,472,286
Top Prize	698,385	1,231,003	500	5,873,452
Total Combinations	9,937,384	18,912,271	1000	54,979,155
Percent age 45-65	17.233	2.154	13.663	23.663
Percent nonwhite	15.537	13.012	1.421	70.346
Percent college	19.941	5.379	2.1	33.3
Percent urban	69.576	15.632	32.2	100
Per Capita Income	23,292.9	3,594.1	17,714	33,435
Per Capita Lottery Retail Outlets	0.00089	0.00224	0.00058	0.00167

Note: Based on full sample size of 135

Table 2 - Game Characteristic as Predictors of Lottery Sales

Variable	All Lottery Games		Lotto Games		Small Prize Games	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	8.799	4.677	-21.073	-25.312	63.731	136.011
	(0.27)	(0.15)	(0.66)	(1.031)	(0.99)	(1.11)
1/EP	--------	0.309	--------	0.382	--------	-48.823
		(0.11)		(0.17)		(0.80)
Variance[a]	--------	-0.157	--------	-0.126	--------	81.864
		(1.51)		(1.03)		(0.45)
Top Prize[a]	--------	0.149***	--------	0.158***	--------	-74.560
		(3.12)		(2.91)		(0.55)
Total Combinations[a]	--------	-0.0056**	--------	-0.0051*	--------	4.656
		(2.33)		(1.84)		(0.13)
Percent age 45-65	1.449	1.106	0.872	-0.0088	1.612	1.830
	(1.32)	(1.11)	(0.69)	(0.01)	(1.11)	(1.16)
Percent nonwhite	0.807**	0.755**	0.278	0.139	1.495***	1.428***
	(2.24)	(2.05)	(0.80)	(0.50)	(3.54)	(4.50)
Percent college	0.234	0.258	-0.0068	0.132	0.315	0.319
	(0.75)	(0.91)	(0.02)	(0.45)	(0.82)	(0.55)
Percent Urban	-0.074	-0.190	0.194	0.0090	-0.415	-0.278
	(0.34)	(0.95)	(0.78)	(0.05)	(1.12)	(0.72)
Per capita Income	0.0021***	0.0026***	0.0012	0.0018***	0.0031***	0.0033**
	(3.13)	(4.78)	(1.29)	(2.89)	(2.74)	(2.56)
Per capita lottery outlets	7967.1	5330.1	9605.9	4758.7	7162.9	15269.9
	(0.89)	(0.63)	(0.92)	(0.61)	(0.45)	(0.95)
Adjusted R^2	0.367	0.445	0.154	0.521	0.605	0.607
F-test for H_o: all game slopes = 0	--------	5.44***	--------	16.508***	--------	1.062
Breusch-Pagan χ^2	112.79***	172.00***	38.06***	51.99***	12.20*	11.91
Hausman Endogeneity Test (χ^2) - Top Prize	-------	0.643	--------	0.023	--------	0.432
Sample Size	135	135	88	88	47	47

Note: *** denotes significance at 1%, ** at 5%, and * at 10%. Dependent variable is per capita lottery sales. Absolute *t*-statistics in parentheses. White's standard errors used for models (1) - (5).
[a] The coefficients are interpreted as per a 10,000 unit change in the top prize, total combinations, and variance.

Table 3 - Out of Sample Power of the Lotto Model

	Actual Per Capita PowerBall Sales	Estimated Per Capita PowerBall Sales	Forecast Error
Prior to change in prize structure	$25.17	$27.16	+ $1.99
After change in prize structure	$30.67	$29.76	- $0.91

Note: Estimates calculated using coefficients from the Lotto regression model shown in column (4) of Table 2.

References

Alm, James, Michael McKee, and Mark Skidmore. "Fiscal Pressure, Tax Competition, and the Introduction of State Lotteries." *National Tax Journal,* vol. 46, no. 4 (December 1993): 463-476.

Clotfelter, Charles T. "On the Regressivity of State-Operated 'Numbers' Games." *National Tax Journal*, vol. 32, no. 4 (December 1979): 543-548.

Clotfelter, Charles T. and Philip J. Cook. "Implicit Taxation in Lottery Finance." *National Tax Journal*, vol. 40, no. 4 (December 1987): 533-546.

Clotfelter, Charles T. and Philip J. Cook. *Selling Hope: State Lotteries in America.* Harvard University Press, Cambridge, Massachusetts, 1989.

Davis, J. Ronnie, John E. Filer, and David L. Moak. "The Lottery as an Alternative Source of State Revenue." *Atlantic Economic Journal*, vol. 20, no. 2 (June 1992): 1-10.

Deboer, Larry. "Administrative Costs of State Lotteries." *National Tax Journal*, vol. 38, no. 4 (December 1985): 479-87.

Garrett, Thomas A. "A Test of Shirking under Legislative and Citizen Vote: The Case of State Lottery Adoption." *The Journal of Law and Economics,* vol. 42, no. 1 (April 1999): 189-208.

Garrett, Thomas A. "The Leviathan Lottery? Testing the Revenue Maximization Objective of State Lottery Agencies as Evidence for Leviathan." *Public Choice*, 2002.

Godfrey, L.G. *Misspecification Tests in Econometrics: The Lagrange Multiplier Principle and Other Approaches*. Cambridge: Cambridge University Press, 1988.

Gulley, O. David and Frank A. Scott. "The Demand for Wagering on State-Operated Lotto Games." *National Tax Journal*, vol. 45, no. 1 (March 1993): 13-22.

Hansen, Ann. "The Tax Incidence of the Colorado State Instant Lottery Game." *Public Finance Quarterly*, vol. 23, no. 3 (July 1995): 385-98.

Hausman, Jerry A. "Specification Tests in Econometrics." *Econometrica,* vol. 46 (November 1978): 1251-1271.

Mason, Paul M., Jeffrey W. Steagall, and Michael M. Fabritius. "The Elasticity of Demand for Lotto Tickets and the Corresponding Welfare Effects." *Public Finance Review*, vol. 25, no. 5 (September 1997): 474-790.

Mikesell, John L. and C. Kurt Zorn. "State Lotteries as Fiscal Savior or Fiscal Fraud: A Look at the Evidence." *Public Administration Review*, vol. 46 (1986): 311-320.

Mikesell, John L. "A Note on the Changing Incidence of State Lottery Finance." *Social Science Quarterly*, vol. 70, no. 2 (June 1989): 513-21.

Public Gaming International. *Public Gaming Magazine*, vol. 24, no. 8 (September 1996).

Scoggins, John F. "The Lotto and Expected Net Revenue." *National Tax Journal*, vol. 48, no. 1 (March 1995): 61-70.

Scott, Frank A. and John Garen. "Probability of Purchase, Amount of Purchase, and the Demographic Incidence of the Lottery Tax." *Journal of Public Economics*, vol. 54, no. 1 (May 1994): 121-143.

Scott, Frank A. and O. David Gulley. "Testing for Efficiency in Lotto Markets." *Economic Inquiry*, vol. 33 (April 1995): 175-188.

Stover, Mark. "Revenue Potential of State Lotteries." *Public Finance Quarterly*, vol. 15, no.4 (October 1987): 428-440.

Stranahan, Harriet A. and Mary O. Borg. "Separating the Decisions of Lottery Expenditures and Participation: A Truncated Tobit Approach." *Public Finance Review*, vol. 26, no. 2 (March 1998): 99-117.

Thiel, Stuart E. "Policy, Participation and Revenue in Washington State Lotto." *National Tax Journal*, vol. 44, no. 2 (June 1991): 225-35.

Vrooman, David H. "An Economic Analysis of the New York State Lottery." *National Tax Journal*, vol. 29 (1976): 482-489.